Towards realistic needle insertion training simulator using partitioned model order reduction

Félix Vanneste[1], Claire Martin[2], Olivier Goury[1], Hadrien Courtecuisse[2], Erik Pernod[3], Stephane Cotin[2], and Christian Duriez[1]

[1] INRIA, Lille, France
[2] INRIA, Strasbourg, France
[3] InfinyTech3D, Colle sur Loup, France, https://infinytech3d.com

Abstract. Needle-based intervention is part of minimally invasive surgery and has the benefit of allowing the reach of deep internal organ structures while limiting trauma. However, reaching good performance requires a skilled practitioner. This paper presents a needle-insertion training simulator for the liver based on the finite element method. One of the main challenges in developing realistic training simulators is to use fine meshes to represent organ deformations accurately while keeping a real-time constraint in the speed of computation to allow interactivity of the simulator. This is especially true for simulating accurately the region of the organs where the needle is inserted. In this paper, we propose the use of model order reduction to allow drastic gains in performance. To simulate accurately the liver which undergoes highly nonlinear local deformation along the needle-insertion path, we propose a new partition method for model order reduction: applied to the liver, we can perform FEM computations on a high-resolution mesh on the part in interaction with the needle while having model reduction elsewhere for greater computational performances. We show the combined methods with an interactive simulation of percutaneous needle-based interventions for tumor biopsy/ablation using patient-based anatomy.

Keywords: FEM Simulation · Model Order Reduction · Percutaneous intervention.

1 Introduction

Needle-based interventions are among the least invasive surgical approaches to access deep internal structures into organs' volumes without damaging surrounding tissues, they are considered as Minimally Invasive Surgery (MIS). Needles only affect a localized area around the needle shaft, reducing this way the occurrence of traumas and risks of postoperative complications [1]. They can be exceedingly complex because the practitioner needs to be accurate in the order of millimeters for the effectiveness of the treatment [2]. To reach such performances, surgeons have to undergo long and arduous training involving lots of practice on animals, cadavers or samples. To reduce this and have a more versatile and effective learning tool, computed-based simulators are created [3, 4].

They particularly fit this need because needle-based MIS adds distance between surgeons and patients so using simulators allows realistic conditions to be reproduced. In addition, interactions with organs are limited and can be rendered with haptic devices. Advanced numerical simulators are now considered as clinically relevant tools for both training and education. To produce high-fidelity simulations and allow for such freedom, physics-based simulation using finite element modeling (FEM) is particularly suited. It allows free and theoretically unlimited user interactions with virtual organs with a lot of variabilities in terms of scenarios to match patient-specific conditions (geometry of organs, accessibility of the tumors, physiological motions of the patient, and pathologies) [5]. FEM has already been used in different medical applications, and validations of their behavior against real organs have already been conducted [6].

Other methods of modeling for needle insertion exist in the literature and are computationally efficient [7, 8] however, clinicians recommended modeling breathing motions of organs surrounding the region of interest for more realism [9], for which FEM is well suited. In [10] coupling between high-rate haptics with FEM was achieved, but computation time remains a challenge with fine meshes.

For all these reasons, creating a precise and real-time simulator for needle insertion in the liver seems, at first, unfeasible. We will show in this paper how we can make a step forward in achieving this goal using model order reduction (MOR) and particularly using a novel approach of partitioned model.

2 Numerical model

2.1 Needle and organs model

The needle is discretised along a line with segments using nodes with 6 degrees of freedom (DOFs) (3 for translations and 3 for rotations). Following standard FEM discretisation, the organs are discretised with a 3D linear tetrahedral mesh using nodes with 3 DOFs (displacement along x, y and z axes). We can now describe the deformation of the organs and the needle using Newton's 2nd law:

$$\mathbf{M}(\mathbf{q})\dot{\mathbf{v}} = \mathbf{P}(t) - \mathbb{F}(\mathbf{q}, \mathbf{v}) + \mathbf{H}^T\boldsymbol{\lambda}, \tag{1}$$

where $\mathbf{q}$ is the vector of positions, $\mathbf{v}$ is the vector of velocities, $\mathbf{M}(\mathbf{q})$ is the mass matrix, $\mathbf{P}$ gathers external forces, $\mathbb{F}$ accounts for internal forces and $\mathbf{H}^T\lambda$ is the vector of constraint forces contributions, which in this case are generated either between each organ interface to model contact force, or between an organ and the needle to model the needle-tissue contact forces. Note that here $\mathbf{q}$ and $\mathbf{v}$ may refer to either the positions and velocities of the needle or those of the organs. The internal forces $\mathbb{F}$ depend on the mechanical behaviour of the organs and the needle and are nonlinear. In this paper, to model the needle, we consider standard elastic beam elements from beam theory [11]. The constitutive law is parameterised with Young Modulus and Poisson's ratio parameters. Similarly, for the organs we consider a linear elastic constitutive law with a co-rotational formulation to account for large rotations.

We perform implicit time integration over the time intervals $[t_0, t_1, ..., t_{n_t}]$, considering $h = t_{n+1} - t_n$, and we linearise internal forces :

$$\mathbb{F}(\mathbf{q_n} + \mathbf{dq}, \mathbf{v_n} + \mathbf{dv}) = \mathbf{f}_{t_n} + \frac{\delta \mathbb{F}}{\delta \mathbf{q}}(t_n)\mathbf{dq} + \frac{\delta \mathbb{F}}{\delta \mathbf{v}}(t_n)\mathbf{dv}. \tag{2}$$

Neglecting viscosity effects ($\frac{\delta \mathbb{F}}{\delta \mathbf{v}}(t_n) = 0$), we obtain:

$$\underbrace{\left(\mathbf{M} + h^2 \frac{\delta \mathbb{F}}{\delta \mathbf{q}}(t_n)\right)}_{\mathbf{A_n}} d\mathbf{v} = \underbrace{-h^2 \frac{\delta \mathbb{F}}{\delta \mathbf{q}} \mathbf{v}_{t_n} - h\left(\mathbf{f}_{t_n} + \mathbf{p}_{t_{n+1}}\right)}_{\mathbf{b_n}} + h\mathbf{H}^T \lambda, \tag{3}$$

with $d\mathbf{v} = \mathbf{v}_{n+1} - \mathbf{v}_n$. Note that $\frac{\delta \mathbb{F}}{\delta \mathbf{q}}(t_n)$ can be identified with the classical tangent stiffness matrix, and $\mathbf{p}_{t_{n+1}}$ is the body force $\mathbf{P}$ at time t_{n+1}.

The simulation is performed by solving this equation at each time step, for each organ and the needle, while respecting the contact constraints. Equation (3) is of large dimension when using fine meshes, which is needed for accurate simulation. This is a major challenge to reach the goal of interactive simulation.

2.2 Standard Model order reduction by projection

In the field of computational mechanics, model order reduction (MOR) aims to represent a fine and accurate numerical model with a surrogate model of small dimension that remains almost as accurate. The position $\mathbf{q}(t)$ is expressed as a truncated expansion of orthonormal vectors :

$$\mathbf{q}\,(t) \approx \mathbf{q}(0) + \sum_{i=1}^{N} \boldsymbol{\phi}_i \alpha_i(t) = \mathbf{q}(0) + \boldsymbol{\Phi}\,\boldsymbol{\alpha}(t), \tag{4}$$

To build the reduced basis $\boldsymbol{\Phi}$ in snapshot-Proper Orthogonal Decomposition (Snapshot-POD), the fine FE model is tested in multiple ways in an offline stage to obtain a database of positions stored in a large snapshot matrix $\mathbf{S}$. A singular value decomposition is performed: $\mathbf{S} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$, with $\mathbf{U}$ and $\mathbf{V}$ unitary matrices and $\boldsymbol{\Sigma}$ a rectangular diagonal matrix containing the singular values of $\mathbf{S}$. The optimal basis of order p is then given by the first p left singular vectors (in $\mathbf{U}$) associated to the p largest singular values σ_i. We can know the truncation error of a POD transform of order p by computing:

$$\nu^2 = \frac{\sum_{i=p+1}^{n_q} \sigma_i^2}{\sum_{i=1}^{n_q} \sigma_i^2}, \tag{5}$$

which is very convenient to guarantee a certain accuracy of the surrogate model. Substituting $\mathbf{q}$ by its reduced expression in equation (3) and projecting the equation onto the reduced basis leads to the reduced equation:

$$\boldsymbol{\Phi}^T \mathbf{A}_n \boldsymbol{\Phi}\, \dot{d\boldsymbol{\alpha}}(t) = \boldsymbol{\Phi}^T \mathbf{b}_n + h(\mathbf{H}\,\boldsymbol{\Phi})^T \boldsymbol{\lambda}. \tag{6}$$

This system is of much smaller dimension than the original fine model and can be solved at much higher rates. However, the integration of the stiffness and mass matrices still have to be performed on the fine mesh before being projected onto the reduced basis. One method to alleviate this issue is to use a method of hyperreduction [12] which allows to perform integration of the mechanical matrices only onto a small subset, called the reduced integration domain (RID) of the FE elements but still save accuracy. In this paper we will use the energy-conserving weighting and sampling method (ECSW) [13]. The reduced equation (6) is further reduced : $\boldsymbol{\Phi}^T \widetilde{\mathbf{A}_n} \boldsymbol{\Phi}\, \dot{d\boldsymbol{\alpha}}(t) = \boldsymbol{\Phi}^T \widetilde{\mathbf{b}_n} + (\mathbf{H}\boldsymbol{\Phi})^T \boldsymbol{\lambda}$, with $\widetilde{\mathbf{A}_n}$ and $\widetilde{\mathbf{b}_n}$ being the restrictions of the original operators to the RID.

2.3 Local Model Order Reduction by projection

The method described above can be used to reduce organs in the direct vicinity of the liver, which influences its boundary conditions. But we can't reduce the liver in the same way: where the needle is inserted, the tissue deforms strongly locally. A reduced model cannot capture these strong local non-linearities [14, 15]. In this section we propose to apply nonetheless a reduction to the liver but only "far" from the needle, where non-linearities remain small and can be accurately captured by a reduced model, while keeping a full fine model in the liver region where the needle may be inserted. A similar idea was proposed in the computational mechanics community [16] to solve damage problems. We assume that we define this region offline, implying that we subdivide the liver DOFs into two distinct sets:

- a set $\mathcal{F}$ that will be solved using the high resolution mesh without reduction. Quantities in that set will be denoted with the subscript $_F$,
- a set $\mathcal{R}$ that will be reduced. The subscript $_R$ will be used.

For the special case of the liver, we may rewrite equation (3) (ignoring iteration subscript n for clarity):

$$\begin{bmatrix} \mathbf{A}_{RR} & \mathbf{A}_{RF} \\ \mathbf{A}_{FR} & \mathbf{A}_{FF} \end{bmatrix} \begin{bmatrix} \mathbf{dv}_R \\ \mathbf{dv}_F \end{bmatrix} = \begin{bmatrix} \mathbf{b}_R \\ \mathbf{b}_F \end{bmatrix}, \tag{7}$$

where $\mathbf{A}_{RR}$ is the restriction of the mechanical matrices to the DOFs in the set $\mathcal{R}$, $\mathbf{A}_{FF}$ to the set $\mathcal{F}$, and $\mathbf{A}_{RF} = \mathbf{A}_{FR}^T$ the mechanical coupling between the two regions defined by sets $\mathcal{R}$ and $\mathcal{F}$, i.e. the coupling between the reduced part of the liver and the part that is solved on the high resolution mesh. In $\mathbf{b}_R$ are included forces due to contact and in $\mathbf{b}_F$ forces due to needle-tissue interactions.

Applying reduction only on the subset r with a basis $\boldsymbol{\phi}$, we obtain:

$$\begin{bmatrix} \boldsymbol{\phi}^T \mathbf{A}_{RR} \boldsymbol{\phi} & \boldsymbol{\phi}^T \mathbf{A}_{RF} \\ \mathbf{A}_{FR} \boldsymbol{\phi} & \mathbf{A}_{FF} \end{bmatrix} \begin{bmatrix} d\boldsymbol{\alpha} \\ \mathbf{dv}_F \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}^T \mathbf{b}_R \\ \mathbf{b}_F \end{bmatrix} \tag{8}$$

We have now a hybrid system that contains both reduced coordinates $\boldsymbol{\alpha}$ and full order FEM coordinates (velocities increment in this case) $\mathbf{dv}_F$. In principle,

we will try to keep the amount of DOFs in the set $\mathcal{F}$ as low as possible to optimise the computational gain since this part is not reduced. Hyperreduction can still be applied on the reduced region, except we force the RID to contain all the elements at the boundary between the 2 regions to keep the coupling intact. We finally obtain the locally reduced equation for the liver:

$$\begin{bmatrix} \boldsymbol{\phi}^T \widetilde{\mathbf{A}_{RR}\boldsymbol{\phi}} & \boldsymbol{\phi}^T \mathbf{A}_{RF} \\ \mathbf{A}_{FR}\boldsymbol{\phi} & \mathbf{A}_{FF} \end{bmatrix} \begin{bmatrix} d\boldsymbol{\alpha} \\ \mathbf{dv}_F \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}^T \widetilde{\mathbf{b}_R} \\ \mathbf{b}_F \end{bmatrix} \tag{9}$$

3 Experimental validation and discussion

We will use the simulation software called SOFA[4] with the open source plugin called MOR[5] where the presented method is implemented. In addition, we will use SOFA GPU implementation based on CUDA.

3.1 Simulation configuration

Anatomy represented The stomach and large intestine, interacting directly with the liver, are accurately represented with a fine mesh. They are fixed to the liver using spring constraints respecting their anatomical attachments. We then add the diaphragm which will emulate the breathing behavior and its repercussions on the displacement of the different organs created by the respiration motion interaction between them. To finish we add the intestine, which will be represented as a coarse hexahedra grid mapped on a collision mesh. This simple model is sufficient due to the weak contribution of the intestine to the general motion/deformation. We don't take into account other factors such as the peritoneum, skin, and blood vessels. We aim to use this simulation for the moment to showcase our new reduction method and show its interest for future integration into a much more detailed and complex simulation to create more advanced simulators. Fig.1 shows the simulation visual with the different organs, the needle and the selected insertion area.

Data-set used The mesh used comes from a clinical real data-set publicly available[6] which has different liver meshes presenting tumors that we will use here as intervention example to train tumor biopsy/ablation. We will study these in different simulation scenarios using two different discretisations:

- *Coarse*: about one thousand tetrahedra, four thousand and five thousand tetrahedra for the large intestine, stomach and liver respectively.
- *Fine*: about nine thousand, nineteen thousand and forty thousand tetrahedra for the large intestine, stomach and liver respectively.

We perform reductions on both and compare their respective performances in precision and computing speed.

[4] https://www.sofa-framework.org/
[5] https://github.com/SofaDefrost/ModelOrderReduction
[6] https://www.ircad.fr/research/data-sets/liver-segmentation-3d-ircadb-01/

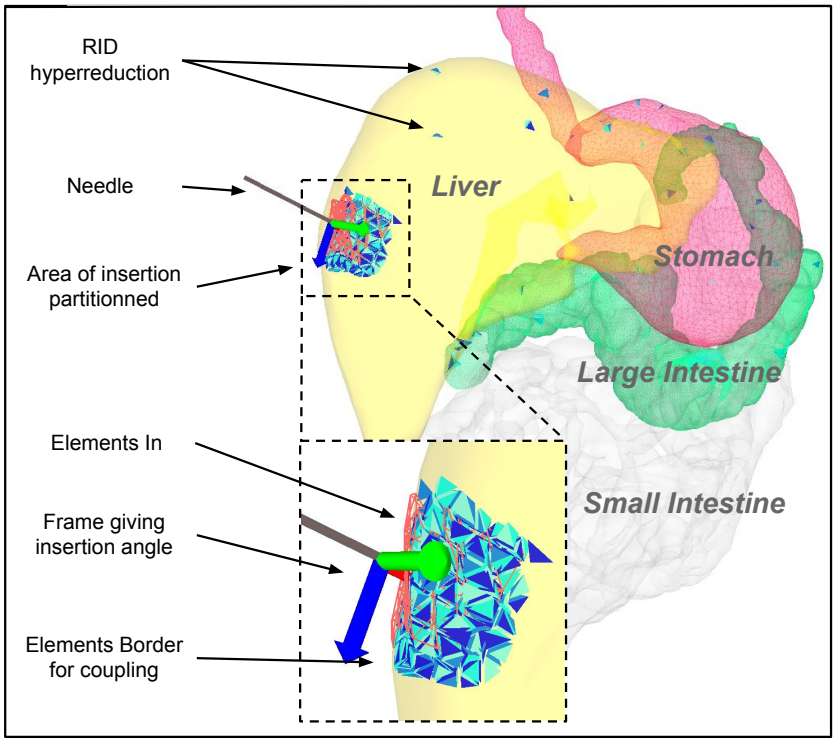

Fig. 1: Simulation setup. Focus on the insertion area where we define the partition of the model with, in orange, the elements strictly inside the set $\mathcal{F}$ defined in section 2.3. This set is enriched with elements on the boundary of the region (in blue) to maintain coupling between the reduced and the fine region.

3.2 Reduced Model generation

We need to stimulate the liver in an offline stage to generate the reduced basis described in section 2.3. To do that we perform a sequence of insertion in a particular area of interest. Here we will simulate reaching a tumor with the needle. Depending on the tumor placement we will choose the corresponding area of insertion, and needle orientation and define a depth to reach the virtual tumor. In addition to this sequence, we perform a little circular motion with the needle when it's fully inserted to stimulate as much as possible nearby elements of the mesh to gather more deformation information. We then remove the needle and put it back to its starting position. This sequence is done at a slow pace to reduce instabilities and inertia, it takes 240 steps. Throughout the whole simulation, the diaphragm emulates breathing to get respiratory motion information into the reduced model. The sequence described is generic and adaptable : we can add new insertion points into a single sequence for different patient data and

Table 1: Computation time and frame rate for each simulation setup based on a fine mesh discretisation.

∼	**CPU only**	**Cuda only**	**Cuda for collisions + MOR**
fps	0.0046	3.7	16.1
ms	216975	240	51

tumor placement, creating a specific reduced model for each scenario. We will display one of these scenario results in the next section.

3.3 Results

Speedup comparisons All the following results were produced using the following PC configuration: CPU: Intel® Core™ i7-7820HQ CPU @ 2.90GHz × 8, GPU: NVIDIA Quadro M1200 Mobile[7], RAM: 16.0Gib. This hardware configuration is old and computation time can easily be outperformed with more up-to-date hardware. However, they give us a good indication and show the general trend of speed-up we can expect. Two reduced models based on the same insertion sequence are tested with the 2 mesh discretisations presented before.

Frame rate (fps) comparisons using different acceleration techniques are displayed in Tab.1.

A completely parallelised version on CUDA for both organ deformation computation and collision solving already brings good speedup. However, the last result clearly shows a significant gain with the presented method using MOR in combination with CUDA for dealing with collision solving. We approach real-time performances with 16.1 fps which is 4.35 times faster than a total CUDA version, we can even expect to reach real-time with a more powerful setup.

Precision comparisons To compare the different models' accuracy, we use the positions from the full-order simulation with the fine mesh presented in the data-set (i.e. with no reduction) as ground truth. We then compute a general mean distance in positions of each simulation compared to it, this will be the error that we want to minimize. We can see the evolution of these means in Fig.2, see supplementary material for an organ per organ position error. We have done 2 comparisons :

- With only the breathing motion to make sure that it was well captured by the reduced model. In Fig.2.A we see that the fine reduced models ($F_{\mathbf{R}}$) are more precise than the coarse models (C) and that the error is much smaller and more constant over time compared to C and the coarse reduced models ($C_{\mathbf{R}}$). The error decrease obtained by the reduced models with fine meshes compared to the full-order coarse models is 50% globally and up to 71% for the liver only. We can also note that, for this simple scenario, the $C_{\mathbf{R}}$ can be used instead of the C because their errors are nearly the same.

[7] https://www.videocardbenchmark.net/gpu.php?gpu=Quadro+M1200&id=3651

- With a needle being inserted in the liver. Results are displayed in Fig.2.B. This time the $C_{\mathbf{R}}$ are less accurate and are not a good alternative to the C. The $F_{\mathbf{R}}$ remains more precise by about 32% compared to the C. The error mostly comes from the large intestine and stomach since we still have the same order of error decrease for the liver only: 65%.

From these two comparisons, if we consider $F_{\mathbf{R}}$ and only the liver where we want to have as little deviation as possible from reality, we obtain an overall error decrease of 68% for a speed-up of 4.35. These results can then be refined according to the needs of the application thanks to the flexibility of model reduction. Indeed we can choose to use more deformation modes inducing a bigger reduced-basis and a slower resolution time but obtaining better precision, or to reduce them and gain speed at the cost of precision. These configurations can easily be adjusted afterwards, showing another powerful point of MOR approach.

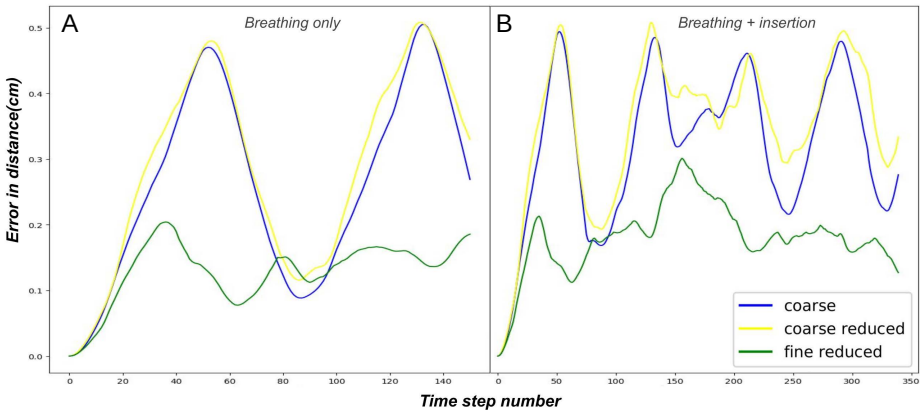

Fig. 2: Aggregate mean distance of all the positions of the simulation to ground truth along the simulation time step for 2 scenarios: **A** breathing only during 150 steps | **B** breathing motion with needle insertion during 340 steps.

4 Conclusion

In this paper, we have presented a new approach to model order reduction (MOR) in the context of real-time soft tissue simulation. We propose reduction by model partitioning, enabling a model to be reduced globally while retaining exact accuracy over a specific area of interest, thus further increasing speed. We have demonstrated the use of the method in a needle-based MIS simulation of the liver. The liver is simulated along with the surrounding organs thanks to MOR acceleration, and we can reproduce the effect of breathing. We believe this is a step forward for real-time, accurate and interactive simulation, with the ultimate goal of creating training simulators for needle-based MIS. The proposed

model can be further improved by selecting only those elements located close to the needle passage. In addition, we can consider placing all problem solving on CUDA to make the most of the GPU's capabilities and avoid copying data between CPU and GPU. We can also investigate the creation of a reduced model that can be activated according to the area of the liver the trainee is working on: for example, by using Couinaud classification to create a reduced model for each of the liver segments and activating/deactivating them accordingly. Further work will continue to add complexity to the model presented, such as the needle/skin interaction and we will also investigate the integration of haptic feedback with an external physical device and its interaction with the simulation, which is an essential step in creating a simulator for trainees.

Acknowledgments. This work was funded by the French National Research Agency (ANR) through the project ANR-21-CE45-0031 SPECULAR.

Disclosure of Interests. The authors have no competing interests.

References

1. N. J. Cowan, K. Goldberg, G. S. Chirikjian, G. Fichtinger, R. Alterovitz, K. B. Reed, V. Kallem, W. Park, S. Misra, and A. M. Okamura, "Robotic needle steering: Design, modeling, planning, and image guidance," Surg. Robot., pp. 557–582, 2011.
2. T. L. De Jong, N. J. van de Berg, L. Tas, A. Moelker, J. Dankelman, and J. J. van den Dobbelsteen, "Needle placement errors: Do we need steerable needles in interventional radiology?," Med. Devices Evid. Res., vol. 11, pp. 259–265, 2018.
3. R. Taschereau, J. Pouliot, J. Roy, and D. Tremblay, "Seed misplacement and stabilizing needles in transperineal permanent prostate implants," Radiother. Oncol., vol. 55, no. 1, pp. 59–63, Apr. 2000.
4. J. T. Hing, A. D. Brooks, and J. P. Desai, "Reality-based needle insertion simulation for haptic feedback in prostate brachytherapy," in Proceedings - IEEE International Conference on Robotics and Automation, 2006, vol. 2006, pp. 619–624.
5. G. Ravali and M. Manivannan, "Haptic Feedback in Needle Insertion Modeling and Simulation," IEEE Rev. Biomed. Eng., vol. 10, no. c, pp. 63–77, 2017.
6. S. Marchesseau, T. Heimann, S. Chatelin, R. Willinger, and H. Delingette, "Multiplicative Jacobian Energy Decomposition Method for Fast Porous Visco-Hyperelastic Soft Tissue Model," in Lecture notes in computer science, vol. 6361, no. PART 1, Springer, 2010, pp. 235–242.
7. Ma Alamilla, Charles Barnouin, Richard Moreau, Florence Zara, Fabrice Jaillet, et al.. A Virtual Reality and haptic simulator for ultrasound-guided needle insertion. IEEE Transactions on Medical Robotics and Bionics, IEEE, 2022, 10.1109/TMRB.2022.3175095 . hal-03657576
8. Pepley, D. F., Yovanoff, M. A., Mirkin, K. A., Miller, S. R., Han, D. C., Moore, J. Z. (2018). "Integrating cadaver needle forces into a haptic robotic simulator". Journal of Medical Devices, Transactions of the ASME, 12(1). https://doi.org/10.1115/1.4038562
9. Di Vece, C., Luciano, C., & de Momi, E. (2021). "Psychomotor skills development for Veress needle placement using a virtual reality and haptics-based simulator". International Journal of Computer Assisted Radiology and Surgery, 16(4), 639–647. https://doi.org/10.1007/s11548-021-02341-0

10. Peterlik, I., Duriez, C., Cotin, S. (2011). "Asynchronous Haptic Simulation of Contacting Deformable Objects with Variable Stiffness". https://doi.org/10.0/Linux-x86_64
11. Przemieniecki, Janusz S. Theory of matrix structural analysis. Courier Corporation, 1985.
12. Ryckelynck, D., 2005. A priori hyperreduction method: an adaptive approach. Journal of computational physics, 202(1), pp.346-366.
13. Goury, Olivier, and Christian Duriez. "Fast, generic, and reliable control and simulation of soft robots using model order reduction." IEEE Transactions on Robotics 34.6 (2018): 1565-1576.
14. Kerfriden, P., Goury, O., Rabczuk, T. and Bordas, S.P.A., 2013. A partitioned model order reduction approach to rationalise computational expenses in nonlinear fracture mechanics. Computer methods in applied mechanics and engineering, 256, pp.169-188.
15. Bordas, S., Kerfriden, P., Courtecuisse, H., Duriez, C., Cotin, S., Goury, O., ... Hale, J. (2016). Simulating topological changes in real time for surgical assistance.
16. Kerfriden, P., Passieux, J.C., Bordas, S.P.A. (2012) Local/global model order reduction strategy for the simulation of quasi-brittle fracture.